AD-A096 130 STANFORD UNIV CA SYSTEMS OPTIMIZATION LAB F/G 12/1

EXPECTED NUMBER OF STEPS OF THE SIMPLEX METHOD FOR A LINEAR PRO--ETC(U)

OCT 80 G B DANTZIG N00014-75-C-0267

UNCLASSIFIED SOL-80-3R NL

1 OF 1

AD A096130

END

DATE FILMED

4-81

DTIC

LEVEL

Systems Optimization Laboratory

Department of Operations Research
Stanford University
Stanford, CA 94305

SYSTEMS OPTIMIZATION LABORATORY
DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305

See 1473.

EXPECTED NUMBER OF STEPS OF THE SIMPLEX METHOD FOR A LINEAR PROGRAM WITH A CONVEXITY CONSTRAINT

by

George B. Dantzig

TECHNICAL REPORT SOL 80-3R

REVISED October 1980

Research and reproduction of this report were supported by the Department of Energy Contract DE-AC03-76SF00326, PA# DE-AT03-76ER72018; Office of Naval Research Contract N00014-75-C-0267; National Science Foundation Grants MCS76-81259, MCS-7926009 and ECS-8012974 (formerly ENG77-06761).

EXPECTED NUMBER OF STEPS OF THE SIMPLEX METHOD FOR A LINEAR PROGRAM WITH A CONVEXITY CONSTRAINT

by

GEORGE B. DANTZIG

Abstract

When there is a convexity constraint, $\Sigma\lambda_j = 1$, *each iteration* t *of the simplex method provides a value* z_t *for the objective and also a lower bound* $z_t - w_t$. *The paper studies (1) the expected behavior of* (w_t/w_0), *(2) probability of termination on the* t-*th iteration, and (3) the expected number of steps,* ξ *ITER, when the columns are drawn from a distribution from a three parameter class of distributions. Using estimates based on a random like behavior for covering simplices, it is shown that*

$$\xi\, ITER \leq m[2.8 + \log_e(\theta_1\theta_2) + (2.8/f)\log_e \bar{n}], \qquad \theta_i \geq 1,$$

where $n = \bar{n} + m + 1$ *is the number of non-negative variables,* $m+1$ *the number of equations.* θ_i *and* f *are parameters for varying the distribution.*

Reasonable bounds are $1.5 \leq \theta_1\theta_2 \leq 4$. *The critical parameter is* $f > 0$. *Poor performance can be expected if* $f << 1$. *For* $\theta_1\theta_2 = 4$, *and*

$$f = 1: \qquad \xi ITER \leq 4.2m + 2.8m \log \bar{n},$$

$$f = m/2: \qquad \xi ITER \leq 4.2m + 5.5 \log \bar{n}.$$

It is conjectured that $f = m/2$ *may be typical of practical problems. If so, for large* m *and* $\bar{n} \leq$ *some fixed multiple of* m, ξ *ITER* $< 4.2\,m$ *iterations as* $m \to \infty$. *Tighter bounds for* $m \leq 5000, n \leq 4m$ *are tabulated. For* $m = 1000, n \leq 4000$, *and* $f = m/2$, $\xi ITER < 1.5\,m$.

The Approach

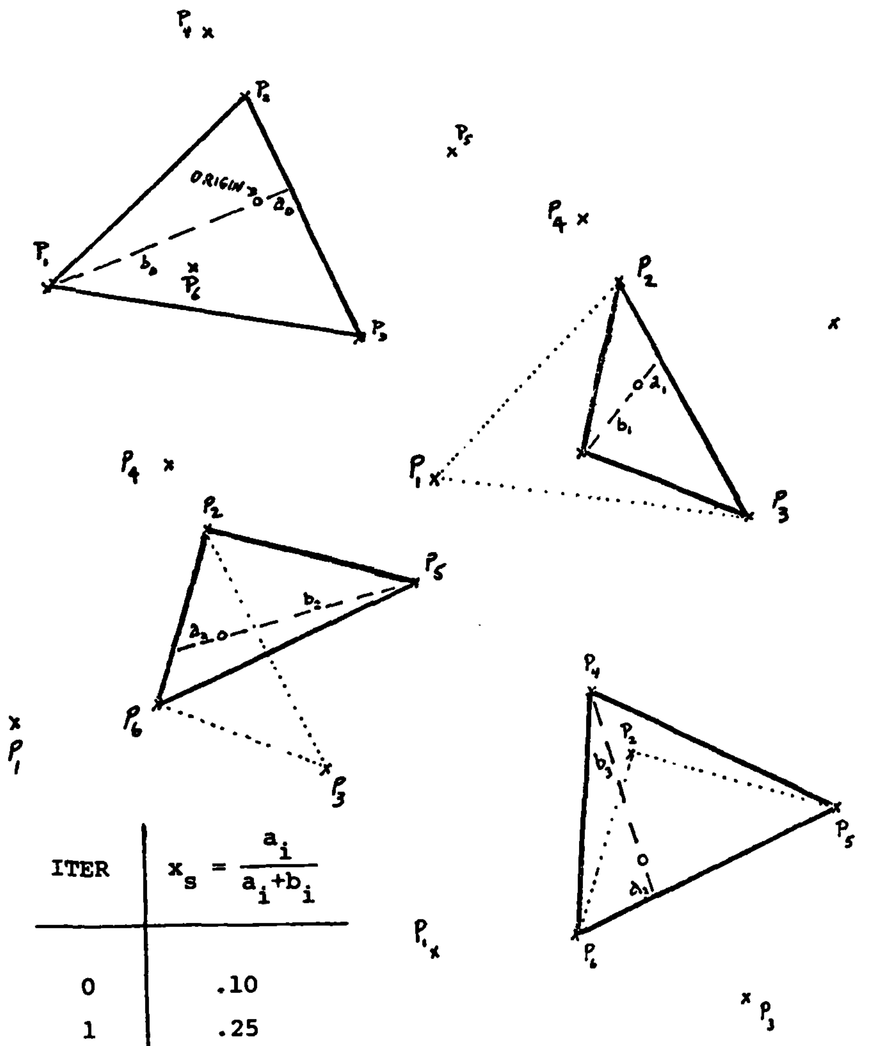
Instead of viewing linear programs as a single class of problems and then applying a worst case analysis, I have taken the view that they should be classified by the characteristics of the distribution from which the columns are drawn. A special class of distributions with three parameters θ_0, $\theta_1\theta_2$, f is studied. A bound on the expected number of steps as a function of these parameters is obtained. Given a linear program, one could use its input data to estimate values for θ_i and f and then use the formulae to bound the expected number of steps. Much work remains to find a good way to characterize linear programs encountered in practice and to develop formulae for estimating the number of steps as a function of their parameters.

Each step of the simplex method with a convexity constraint produces in the space of the columns a simplex that covers a point $\bar{b}$, corresponding to the right hand side. The process that locates $\bar{b}$ relative to the simplex, is viewed as a kind of black box out of which pops the random value of the incoming variable λ_s. If the point $\bar{b}$ (as expressed by its barycentric coordinates in the simplex) is uniformly distributed in the simplex, then $\rho = 1 - \lambda_s$ has the density distribution $m\,\rho^{m-1}d\rho$, $0 \leq \rho \leq 1$ and expected value $\lambda_s = 1/(m+1)$.

The idea is illustrated in Figure 0. Starting at top left, the simplex $P_1P_2P_3$ covers the origin with P_2 the entering column. The value of the incoming variable λ_s is $a_0/(a_0 + b_0)$. Next, top right, P_6 replaces P_1, yielding $\lambda_s = a_1/(a_1 + b_1)$. Next, at bottom left, P_5 replaces P_3. Then, at bottom right, P_4 replaces P_2. Note the "random" behavior of the origin's location relative to various simplices, particularly the value of λ_s.

By making estimates based on $m\rho^{m-1}d\rho$, we can bypass the difficult (if not intractable) analysis of the number of edges (steps) along the path of edges in the polyhederal set generated by the simplex method. I believe this distribution leads

to a higher estimate of the number of steps than the true one but much work remains to show it is so or to find a better one to take its place.



ITER	$x_s = \frac{a_i}{a_i + b_i}$
0	.10
1	.25
2	.20
3	.16

Geometry of the Simplex Algorithm

The problem is to estimate the number of iterations to solve by the simplex method the linear program: FIND $\lambda_j \geq 0$, min z

$$\sum_{j=1}^{n} P_j\lambda_j = 0, \quad \sum_{j=1}^{n} \lambda_j = 1, \quad \sum_{j=1}^{n} c_j\lambda_j = z$$

where P_j are m-vectors. Because of the convexity constraint a problem with a general right hand side $\sum_{j=1}^{n} \overline{P}_j\lambda_j = \bar{b}$ may be reduced to the above by setting $\overline{P}_j = P_j + \bar{b}$. A general linear program, i.e., one without a convexity constraint: FIND $x_j \geq 0$, min z :

$$\sum_{j=1}^{n} \overline{P}_j x_j = \bar{b}, \quad \sum_{j=1}^{n} c_j x_j = z,$$

can be reduced to the above providing one is willing to impose an upper bound M on the sum of the variables

$$\sum_{j=1}^{n} x_j + x_{n+1} = M.$$

This is done by setting $x_j = M\lambda_j$,dividing each equation by M, and then setting $\overline{P}_j = P_j + \bar{b}/M$ as above.

Setting aside the objective function z for the moment, the vectors P_j may be thought of as points scattered around the origin in R^m and we are seeking weights $\lambda_j \geq 0$ to assign to the points so that their *center of gravity* is the origin. See Figure 1. Any $m+1$ points P_{j_i} whose simplex is full dimensional (shaded area) and covers the origin corresponds to a *basic feasible* solution. [In a different geometry namely the space of $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ such a solution corresponds to an *extreme point*.] Any other such simplex could qualify as the optimal basic feasible solution if its corresponding "cost" coefficients c_{j_i} are sufficiently small relative to the other c_j.

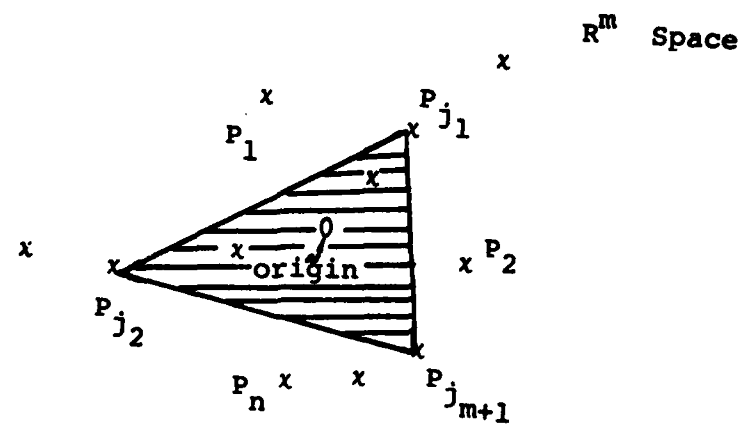


FIGURE 1.

The simplex method has two phases. Both phases use the same procedure but on different problems. Phase I's purpose is to find a *basic feasible* solution to start Phase II. Phase I is set up in such a way that for its problem a starting basic feasible solution is at hand without any computational effort. The two phases use the same algorithm. We will therefore, discuss only the effort to solve the Phase II problem.

The simplex method's Phase II is initiated by a selection (found in Phase I) of $(m+1)$ points P_{j_i} with two properties: *first* that

$$B = \begin{bmatrix} P_{j_1} & P_{j_2} & \cdots & P_{j_{m+1}} \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

is non-singular and *second* the solution to

$$\sum_i P_{j_i} \lambda^0_{j_i} = 0 \quad , \qquad \qquad \sum_i \lambda^0_{j_i} = 1 \quad ,$$

yields $\lambda^0_{j_i} \geq 0$. The columns of B therefore form a *basis* in the space of the columns $(P_j, 1)^T$.

The basic column indices are denoted by

$$B = \{j_1, j_2, \dots, j_{m+1}\} \quad ,$$

and the solution $\lambda_{j_i} = \lambda_{j_i}^0 \geq 0$ for $j \in B$ and $\lambda_j = 0$ otherwise, is called a *basic feasible solution*. The *value* of this solution is

$$z = z_0 = \sum_{i=1}^{m+1} c_{j_i} \lambda_{j_i}^0 \quad .$$

The iterative step consists of replacing a column P_{j_r} by P_s in such a way that the two properties above are preserved and there is a decrease in the value of z. If the origin is in the *interior* of the simplex, the decrease is *strict*. If this happens on each step, it is easy to see, since there are only a finite number of simplices, that the process is finite. However, if the origin is on a lower dimensional face of the simplex and the incoming point P_s does not replace a point on this face, there will be no improvement. In such a case it is necessary to use a perturbation scheme to get around the "degeneracy" in order to guarantee convergence in a finite number of steps. In our approach the columns of the linear program are selected at random from a distribution. For the class of distributions studied, the probability of the origin lying on a face of a covering simplex, is zero and therefore degeneracy need not be considered.

Each step (iteration) computes (Π, π_{m+1}),

$$\Pi P_{j_i} + \pi_{m+1} = c_{j_i} \quad , \qquad j_i \in B \quad .$$

This system is solved using B^{-1} which is updated by $(m+1)^2$ multiplications and additions. Letting

$$\delta_j = c_j - \Pi P_j - \pi_{m+1} \quad ,$$

the incoming column P_s is selected by

$$s = \operatorname{argmin} \delta_j \quad .$$

Note that

$$\delta_s \leq \delta_j \qquad \text{and} \qquad \delta_{j_i} = 0\,, \qquad j_i \in B \quad .$$

It is easy to see that for all feasible solutions $z_0 + \delta_s$ is a *lower bound* for z_1. Indeed if $\delta_s \geq 0$, then the current solution z_0 is optimal and the iterative process stops. To prove $z_0 + \delta_s$ is a lower bound:

$$0 = \sum_{i=1}^{m+1} \delta_{j_i}\lambda^0_{j_i} = \sum_{i=1}^{m+1} (c_{j_i} - \Pi P_{j_i} - \pi_{m+1})\lambda^0_{j_i} = \sum_{i=1}^{m+1} c_{j_i}\lambda^0_{j_i} - \pi_{m+1} = z_0 - \pi_{m+1}$$

$$\sum_{j=1}^{n} \delta_j\lambda_j = \sum_{j=1}^{n}(c_j - \Pi P_j - \pi_{m+1})\lambda_j = \sum_{j=1}^{n} c_j\lambda_j - \pi_{m+1} = z - z_0 \quad .$$

Hence

$$z - z_0 = \sum \delta_j\lambda_j \geq \min \delta_j = \delta_s \quad , \qquad \lambda_j \geq 0\,, \Sigma\lambda_j = 1\,.$$

Instead of plotting $P_j \in R^m$, we now plot $(P_j, c_j) \in R^{m+1}$, see Figure 2. The problem is to find weights $\lambda_j \geq 0$ so that the center of gravity lies on the line $(0, 0, \ldots, z)$, which we will refer to as the *z axis*, and such that the z coordinate is minimum. The hyperplane, called the *solution plane*, that passes through the points (P_{j_i}, c_{j_i}) associated with the basis B, has for equation

$$\Pi P + \pi_{m+1} = z \quad , \qquad \pi_{m+1} = z_0$$

where (P, z) is in the space of all possible columns (P_j, c_j). This plane intersects the z axis at G with $z = z_0$. The point $L = (P_s, c_s)$ selected for improvement, is the point (P_j, c_j) whose absolute distance below the hyperplane is the greatest.

The *lower bound* point F on the z axis with $z = z_0 + \delta_s = d_t$ is obtained by passing a plane parallel to $\Pi P + z_0 = z$ through $L = (P_s, c_s)$ and finding

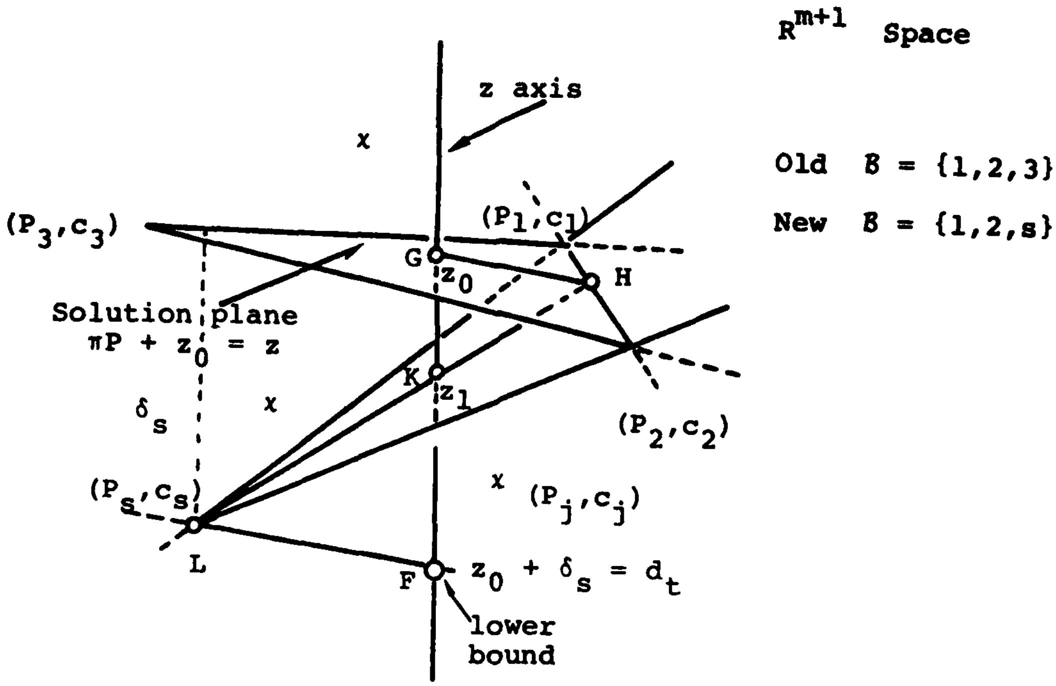


FIGURE 2:

Improving a Basic Solution

where it cuts the z axis. Next find the new solution plane obtained by replacing (P_{j_r}, c_{j_r}) by $L = (P_s, c_s)$ and let it intersect the z axis at K at $z = z_1$. Join L to K and extend until it hits the old solution plane at H. The lines GH and LF are parallel because they lie in the same plane and at the same time they are on two parallel planes. Therefore,

$$\frac{GK}{KF} = \frac{HK}{KL}$$

where $HK/(HK + KL) =$ value of incoming variable λ_s.

We now project parallel to z axis the points (P_j, c_j) onto the $z = 0$ plane and denote by $\bar{L}, \bar{K}, \bar{H}$ the points corresponding to L, K, H. The point K projects into $\bar{K}$, the origin, see Figure 3. Note that

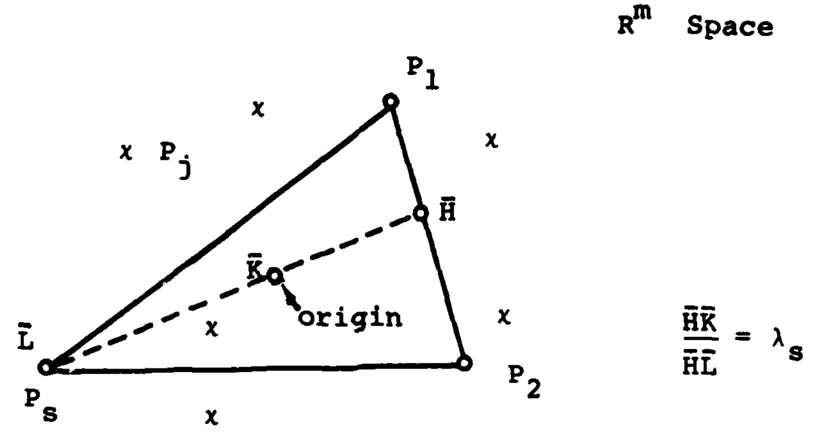


FIGURE 3:

Projection onto $z = 0$

$$\frac{GK}{GF} = \frac{HK}{HL} = \frac{\overline{HK}}{\overline{HL}} = \lambda_s \quad .$$

The smaller $GF = (-\delta_s)$ and λ_s, the less the improvement GK because

$$GK = \lambda_s(-\delta_s).$$

The basic solution $(\lambda_{j_1}, \lambda_{j_2}, \ldots, \lambda_{j_{m+1}})$ is also the barycentric coordinates of $\overline{K}$ in the simplex for iteration r. If P_s just entered replacing P_{j_r} then P_s corresponds to the r-th coordinate and $\lambda_{j_r} = \lambda_s$, the value of the incoming variable.

Estimating the error w_t after t iterations

On iteration t let z_t be the value of the $t-th$ basic feasible solution and let $z_t - w_t$ be the best lower bound obtained so far. Thus

$$z_t \geq \min z \geq z_t - w_t \quad .$$

Initially for $t = 0$, we have given z_0 and $z_0 - w_0 = \min c_j$. On iteration t, we generate a $\delta_s = \delta_s^t, \lambda_s = \lambda_s^t$. Then

$$z_{t+1} - w_{t+1} = \max(z_t - w_t,\ z_t + \delta_s^t) \quad ,$$

becomes the new best lower bound generated so far. Therefore rearranging

$$\begin{aligned} w_{t+1} &= z_{t+1} - \max(z_t - w_t,\ z_t + \delta_s^t) \\ &= z_t + \lambda_s^t \delta_s - \max(z_t - w_t,\ z_t + \delta_s^t) \\ &= -\lambda_s^t(-\delta_s^t) + \min(w_t,\ -\delta_s^t) \\ &\leq w_t(1 - \lambda_s^t) \quad , \end{aligned}$$

where λ_s^t is the value of the incoming variable in the basic set of iteration $t+1$.

To see that the last step holds, note that (1) if $(-\delta_s^t) > w_t$, then the right hand side increases as $(-\delta_s^t)$ decreases towards (w_t); or (2) if $(-\delta_s^t) < w_t$, then the right hand side is $(-\delta_s)(1 - \lambda_s^t)$ and it increases as $(-\delta_s^t)$ increases towards w_t. Thus the maximum is attained at $(-\delta_s) = w_t$. It follows that

$$w_{t+k} \leq w_t(1 - \lambda_s^t)(1 - \lambda_s^{t+1}) \cdots (1 - \lambda_s^{t+k-1}) \quad .$$

Let

$$\rho = 1 - \lambda_{j_r} ,$$

then

$$w_r \leq \rho w_{r-1} = (1 - \lambda_{j_r}) w_{r-1} \,.$$

If now we assume with equal probability that the incoming column P_s corresponds to any vertex i in the simplex, then, letting $\xi(x)$ stand for expected value x and $\xi(x \mid y)$ for expected value of x given y,

$$\xi\rho = \xi(1 - \lambda_i) = 1 - \frac{1}{m+1} \sum_{i=1}^{m+1} \lambda_i = \frac{m}{m+1} , \qquad \Sigma \lambda_i = 1 ,$$

independent of the distribution of λ_t in the simplex. Therefore

$$\begin{aligned}
\xi(w_t \mid w_{t-1}) &\le [m/(m+1)]w_{t-1}\,; \\
\xi(w_t \mid w_{t-2}) &= \xi[\exp(w_t/w_{t-1} \mid w_{t-1})w_{t-1} \mid w_{t-2}] \\
&\le [m/(m+1)]\xi[w_{t-1} \mid w_{t-2}] \le [m/(m+1)]^2 w_{t-1}\,.
\end{aligned}$$

It follows inductively

$$\xi w_t \le [m/(m+1)]^t w_0 \doteq e^{-t/m} w_0\,.$$

Estimating $\Delta_t = (z_t - d_t) = -\delta^t_s$ after t iterations

As developed earlier, except now Δ_t in place of $-\delta^t_s$:

$$w_t \ge z_t - z_{t+1} = \lambda^t_s \Delta_t\,.$$

Therefore,

$$\Delta_t \le w_t \lambda_s^{-1}\,, \qquad \lambda_s = \lambda^t_s\,.$$

In estimating the number of steps we set aside those steps in which $\lambda_s < \lambda^*_s$ where $\lambda^*_s = \mu/(m+1)$ for some $\mu < 1$ to be chosen later. Steps with $\lambda_s < \lambda^*_s$ may be thought of as "almost" degenerate pivots. We conservatively estimate for them a *zero* improvement even though in fact there is some. Let

$$\beta = \Pr(\lambda_s < \lambda^*_s)\,.$$

To estimate β as a function of μ, **we assume the location of $\overline{K}$ in the simplex is uniform over the simplex.** If so the density distribution of $\rho = (1 - \lambda_s)$ is $m\rho^{m-1}d\rho$. This randomization assumption is consistent with the earlier one in that $\xi\,\rho = \xi(1 - \lambda_s) = m/(m+1)$ as before. We now have

$$\beta = \Pr(\lambda_1 < \lambda^*_s) = \int_{1-\lambda^*_s}^{1} m\,\rho^{m-1}\,d\rho = 1 - (1 - \lambda^*_s)^m\,,$$

$$\begin{aligned}
1 - \beta &= (1 - \lambda^*_s)^m = (1 - \frac{\mu}{m+1})^m \\
&= \{(1 - \frac{1}{(m+1)/\mu})^{(m+1)/\mu}\}^{\mu m/(m+1)} \doteq e^{-\mu}\,.
\end{aligned}$$

Thus if T is the estimated number of steps with $\lambda_s \geq \lambda_s^* = \mu/(m+1)$, then

$$\gamma T = T/(1-\beta) \doteq e^{\mu} T\,, \qquad \mu \leq 1\,,$$

is the estimated number including $\lambda_s < \lambda_s^*$. What we will do, accordingly, is to bound Δ_t *given* $\lambda_s \geq \lambda_s^*$. It is to be understood that the subscripts for $\Delta_t, \Delta_{t+1}, \ldots$, now refer *only* to those steps with $\lambda_s \geq \lambda_s^*$, skipping over the steps $\lambda_s < \lambda_s^*$. The bounds determined for the expected number of iterations with $\lambda_s \geq \lambda_s^*$ will later be corrected for the omitted iterations with $\lambda_s < \lambda_s^* = \mu/(m+1)$ by multiplying by $\gamma = e^{\mu}$.

Therefore for some $0 \leq \phi_\tau \leq 1$,

$$\Delta_\tau = (1-\phi_\tau) w_\tau / \lambda_s \quad , \qquad \lambda_s \geq \lambda_s^*\,.$$

Earlier we showed, except now $\Delta_\tau = -\delta_s^\tau, \lambda_s = \lambda_s^\tau$:

$$\begin{aligned} w_{\tau+1} &= -\Delta_\tau \lambda_s + \min(w_\tau, \Delta_\tau) \\ &= -(1-\phi_\tau) w_\tau + \min(w_\tau, \Delta_\tau) \leq \phi_\tau w_\tau \\ &\leq \phi_\tau \phi_{\tau-1} \cdots \phi_t w_t \end{aligned}$$

and

$$\Delta_\tau \leq (1-\phi_\tau)\phi_{\tau-1}\phi_{\tau-2}\cdots\phi_t w_t/\lambda_s^*\,, \qquad \tau \geq t\,, \quad 0 \leq \phi_\tau \leq 1\,.$$

We will now show that a high value for Δ_t, i.e., ϕ_t close to 0 implies a low upper bound for Δ_{t+1}. Indeed this is clear by noting:

$$\begin{aligned} \Delta_{t+2} &\leq (1-\phi_{t+2})\phi_{t+1}\phi_t\, w_t/\lambda_s^* \\ \Delta_{t+1} &\leq (1-\phi_{t+1})\phi_t\, w_t/\lambda_s^* \\ \Delta_t &\leq (1-\phi_t) w_t/\lambda_s^*\,. \end{aligned}$$

Accordingly our approach is to estimate a bound for Δ_k/w_k by averaging $\alpha_\tau = \Delta_\tau/w_\tau$ over $\tau = k, k-1, \ldots, t$. Denoting the average by $\overline{\alpha}_k$, we will use the bound for $\overline{\alpha}_k$ as the smoothed bound for α_k. By definition

$$\overline{\alpha}_k = \frac{1}{k-t+1}\sum_{\tau=t}^{k}\frac{\Delta_\tau}{w_\tau}.$$

For example for $k = 2$,

$$\begin{aligned}
\overline{\alpha}_{t+1} &= \frac{1}{2}[\frac{\Delta_{t+1}}{w_{t+1}} + \frac{\Delta_t}{w_t}], && w_t \geq w_{t+1} \\
&\leq \frac{1}{2}[\frac{1}{w_{t+1}}(1-\phi_{t+1})\phi_t + \frac{1}{w_t}(1-\phi_t)](w_t/\lambda_s^*), && 0 \leq \phi_i \leq 1 \\
&\leq \frac{1}{2}\frac{1}{w_{t+1}}(w_t/\lambda_s^*) \quad .
\end{aligned}$$

Note that the unsmoothed bound for α_{t+1} is $1/\lambda_s^*$ whereas the bound smoothed over two iterations is about half. Over $k-t+1$ iterations we have

$$\overline{\alpha}_k \leq \frac{1}{k-t+1}\sum_{\tau=t}^{k}\frac{1}{w_\tau}(1-\phi_\tau)\phi_{\tau-1}\cdots\phi_t(w_t/\lambda_s^*), \qquad w_\tau \geq w_{\tau+1}.$$

The values of ϕ_τ are not known, so we choose $0 \leq \phi_\tau \leq 1$ so as to maximize the right hand side. The terms with indices $\tau = k$ and $\tau = k-1$ only involve ϕ_k and ϕ_{k-1} which for fixed $\phi_{k-2}, \ldots, \phi_t$ are maximized by setting $\phi_k = 0$ and $\phi_{k-1} = 1$. The second term dropping, only the first and third involve ϕ_{k-2} and are maximized by setting $\phi_{k-2} = 1$. Continuing in this manner, we obtain $\phi_\tau = 1$ for $\tau = k-1, \ldots, t$ and

$$\overline{\alpha}_k \leq \frac{1}{k-t+1}\frac{1}{w_k}(w_t/\lambda_s^*) = \frac{w_t/w_k}{(k-t+1)\lambda_s^*}.$$

Since we are free to choose $k-t+1$, the number of α_τ to average. Note that $\xi w_k \leq e^{(t-k)/m} \cdot w_t$. It is natural to think of w_k as being roughly proportional

to $e^{(t-k)/m}$. If so the smallest bound for $\bar{\alpha}_k$ can be found by setting to zero the derivative with respect to t; the min is attained at $t = k - m + 1$. It is more convenient to use $t = k - m$. Thus

$$\bar{\alpha}_k w_k \leq w_{k-m}/[(m+1)\lambda_s^*] = w_{k-m}/\mu, \qquad \mu = (m+1)\lambda_s^*.$$

We will use this bound for $\bar{\alpha}_k$ as the smoothed bound for $\alpha_k = \Delta_k/w_k$. Thus

$$\Delta_k \leq \mu^{-1} w_{k-m}, \qquad k \geq m,$$

when $k \geq m$. For $k < m$, $\bar{\alpha}_k$ is defined as the average value of Δ_r/w_r for $0 \leq r \leq m$ yielding

$$\Delta_k \leq \mu^{-1} w_k(w_0/w_m), \qquad k < m,$$

for k such that $\lambda_s^k > \lambda_s^* = \mu/(m+1)$.

Assumptions About the Class of Distributions of Points (P, c)

We think of (P_j, c_j) as a *random sample* of n points drawn from a continuous distribution of points (P, c) in C, a convex set. Let the z axis intersect C at lowest value $\min z$ and highest value $\max z$ and let d^* be some fixed value in between. For iteration t, the solution plane is $\Pi^t P + z_t = z$. The point (P, c) most below this plane lies on a parallel hyperplane that touches C at L at a distance $\Delta_t > 0$ below it. A parallel hyperplane through L cuts the z axis at $d_t = z_t - \Delta_t$. See Figure 4.

For any hyperplane $\Pi P + d = z$, we denote by $\Pr(\delta_j < 0)$ the probability of finding points (P, c) below the hyperplane. Suppose we move the hyperplane parallel to itself from a value $d = d_t$ where it just touches C from below at L up to $d = d^*$. Our assumption about the distribution is that its cumulative probability for $d_t \leq d \leq d^*$ is bounded as follows:

$$\Pr(\delta_j < 0) \leq \theta_0[(d - d_t)/(d^* - d_t)]^f, \qquad 0 \leq \theta_0 \leq 1, \; f > 0,$$

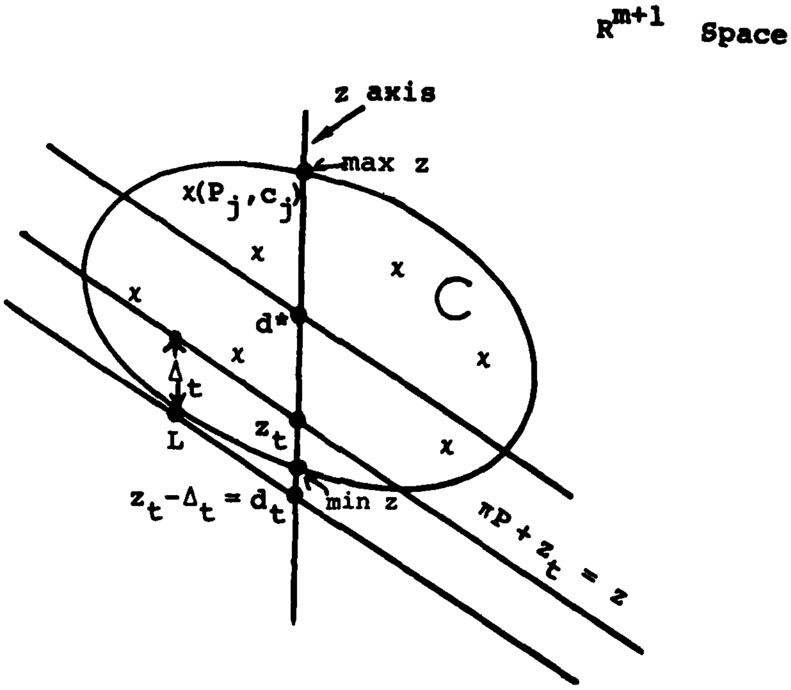


FIGURE 4: $\{(P_j, c_j)\}$ is a sample of n points from a distribution of points (P, c) in C

for some choice of θ_0 and f. If we change the direction Π of the solution plane, d_t will of course change. We are assuming that a fixed θ_0 and f can be found independent of Π.

Case C is a ball

For example, suppose (1) C is a ball in $m+1$ dimensions of unit radius, (2) the distribution of C is uniform in the ball, and (3) the center of the ball is located on the z axis at $z = 1$. Let φ be the angle between the z axis and the normal to the hyperplane, $\overline{\Delta} = \Delta \cos\varphi, \overline{d}^* = 1 - (1 - d^*)\cos\varphi$, and finally let $\overline{\theta}_0 \leq \frac{1}{2}$ be the probability of being in a spherical cap of height $\overline{d}^*$, then for $\overline{\Delta} < \overline{d}^* \leq 1$,

$$\Pr(\delta_j < 0) = \bar{\theta}_0 \int_0^{\bar{\Delta}} (2x - x^2)^{m/2} dx \Big/ \int_0^{\bar{d}^*} (2x - x^2)^{m/2} dx\,, \qquad \bar{\theta}_0 \le \frac{1}{2}\,,$$

$$\le \frac{1}{2} \int_0^{\bar{\Delta}} (2x - x^2)^{m/2} dx \Big/ \int_0^{\bar{d}^*} (2x - x^2)^{m/2} dx\,.$$

If d^* is very small we can ignore the x^2 term and get the inequality

$$\Pr(\delta_j < 0) \le (\Delta/d^*)^{1+m/2}$$

i.e., $f = 1 + m/2$. However, d^* would have to be very small indeed for this to hold. For hyperspheres of dimension $m + 1 \le 21$, I have computed approximate values of the integrals and find that

$$\Pr(\delta_j < 0) \le (\Delta/d^*)^{m/2}\,, \qquad d^* = .4\,.$$

I conjecture for $d^* \le .4$ and $m \to \infty$ that

$$\Pr(\delta_j < 0) \le (\Delta/d^*)^{m/2-k}\,, \qquad k \ge 0\,,$$

where $(k/m) \to 0$. Even if this is not true, it seems intuitively that something like $f = m/2$ should hold. The rough reasoning goes as follows: After m iterations one could safely remove from C all points above say $\Pi^m P + 2z_m = z$ and set $d^* = z_m$. After m iterations, d^* is likely to be sufficiently close to zero that one could then ignore the x^2 terms in the integration as above.

Continuation of general C case

We now reexpress the bound for $\Pr(\delta_j < 0)$ in terms of $w_0, w_t, \Delta_t = z_t - d_t$:

$$\Pr(\delta_j < 0) \le \frac{\theta_0 (z_t - d_t)^f}{[(d^* - z_t) + (z_t - d_t)]^f}\,, \qquad z_t \le d^*\,,$$

$$\le \frac{\theta_0 (z_t - d_t)^f}{[(d^* - \min z - w_t) + (z_t - d_t)]^f}$$

$$\le \theta_0 [1 + (d^* - \min z - w_t)/(z_t - d_t)]^{-f}\,.$$

Referring to Figure 5, we have initially

$$w_0 = z_0 - \min c_j .$$

Let

$$\theta_1 = \frac{\max z - \min z}{d^* - \min z} \quad , \qquad\qquad \theta_1 \geq 1$$

$$\theta_2 = \frac{\max z - \min c_j}{\max z - \min z} \quad , \qquad\qquad \theta_2 \geq 1 .$$

Like f, θ_1 and θ_2 are characteristics of the distribution in C. Reasonable values for θ_1 and θ_2 might be $1.5 \leq \theta_1 \leq 2.5, 1 \leq \theta_2 \leq 2$, see Figure 5. High values of θ_i give rise to high estimates of expected number of steps. For illustrative purposes $\theta_0 = 1$ and $\theta_1\theta_2 = 4$ are used later. Then

$$\theta_2 = \frac{\max z - z_0 + w_0}{\max z - \min z} \geq \frac{w_0}{\max z - \min z} ,$$

or

$$(\max z - \min z) \geq w_0/\theta_2$$

Therefore

$$d^* - \min z = (\max z - \min z)/\theta_1 \geq w_0/\theta_1\theta_2 .$$

It follows for $z_t < d^*$ that,

$$\Pr(\delta_j < 0) \leq \theta_0[1 + \{-w_t + w_0/(\theta_1\,\theta_2)\}/\Delta_t]^{-f} , \qquad \Delta_t = z_t - d_t .$$

In the last section we showed that if we consider only iterations where $\lambda_s \geq \mu/(m+1)$, then Δ_t/w_t had smoothed upper bounds $\mu^{-1}w_{t-m}/w_t$ for $t \geq m$ and $\leq \mu^{-1}w_0/w_m$ for $t < m$. Replacing Δ_t by these upper bounds, we have for t such that $z_t \leq d^*$:

$$\Pr(\delta_j < 0) \leq \theta_0[1 + \mu w_{t-m}^{-1}\{-w_t + w_0/(\theta_1\theta_2)\}]^{-f} , \qquad t \geq m$$

$$\leq \theta_0[1 + \mu w_t^{-1}(w_m/w_0)\{-w_t + w_0/(\theta_1\theta_2)\}]^{-f} , \quad t < m .$$

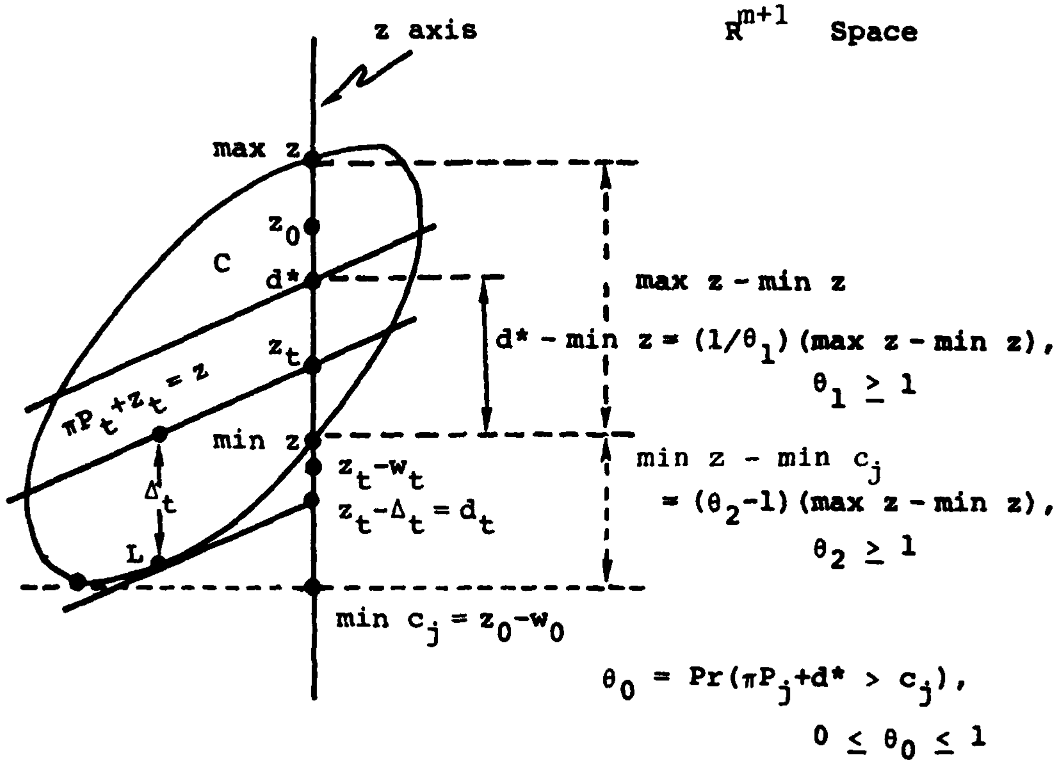


FIGURE 5:

$\theta_0, \theta_1, \theta_2$ are characteristics of the distribution (P, c).

The analysis requires a bound on the ratio Δ_t/w_t.

Estimating $t = t_0$: $z_t \leq d^*$ for $t \geq t_0$.

To estimate t_0, refer to Figure 5. Note that $z_t < d^*$ if $w_t < d^* - \min z$. It therefore holds if

$$w_t \leq w_0/(\theta_1\theta_2)$$

because, as we have just shown, $w_0/(\theta_1\theta_2) \leq d^* - \min z$. Accordingly, we now estimate a $t = t_0$ such that

$$w_{t_0} \leq w_0/(\theta_1\theta_2).$$

Earlier we showed $\xi w_t \leq w_0[m/(m+1)]^t \doteq w_0 e^{-t/m}$. Therefore on the "average"

we overestimate such a t_0 by setting

$$\xi w_{t_0} \leq w_0 e^{-t_0/m} = w_0/(\theta_1\theta_2),$$

$$t_0 = m \log(\theta_1\theta_2), \qquad \theta_i \geq 1,$$

which henceforth will be used as the estimate of the value of t_0 such that $z_t \leq d^*$ for $t \geq t_0$. For iterations $t \geq t_0$, we first determine bounds for expected iterations for t such that $\lambda_s \geq \mu/(m+1)$ and then we multiply these bounds by $\gamma = e^{\mu}$ to correct for omitted iterations having $\lambda_s < \mu/(m+1)$.

Probability of Termination

Given n points (P_j, c_j) in C of which $m+1$ lie on hyperplane $\Pi P + z_t = z$, the probability that none of the remaining $\bar{n} = n-(m+1)$ lie below the hyperplane, i.e., the *probability of termination* on iteration t, given non-termination prior to t, denoted Pr (term), satisfies

$$\Pr(\text{term}) = [1 - \Pr(\delta_j < 0)]^{\bar{n}} \geq p_t, \qquad \bar{n} = n - m - 1$$

where we have just shown

$$p_t = \{1 - \theta_0 [1 + \mu\{-w_t/w_{t-m} + w_0/(w_{t-m}\theta_1\theta_2)\}]^{-f}\}^{\bar{n}}, \qquad t \geq m$$

$$p_t = \{1 - \theta_0[1 + \mu\{-w_m/w_0 + w_m/(w_t\theta_1\theta_2)\}]^{-f}\}^{\bar{n}}, \qquad t < m,$$

for $t \geq t_0$. For $0 \leq t < t_0$, we set $p_t = 0$. We now solve for w_0/w_{t-m} if $t \geq m$ (or w_m/w_t if $t < m$) in terms of $p = p_t$:

$$\frac{1}{\theta_1\theta_2}\frac{w_0}{w_{t-m}} = \frac{1}{\mu}[\theta_0^{1/f}/(1-p^{1/\bar{n}})^{1/f} - 1] + \frac{w_t}{w_{t-m}}, \qquad t \geq m$$

$$\frac{1}{\theta_1\theta_2}\frac{w_m}{w_t} = \frac{1}{\mu}[\theta_0^{1/f}/(1-p^{1/\bar{n}})^{1/f} - 1] + \frac{w_m}{w_0}, \qquad t < m.$$

The value of t given p will be estimated by trying to find a $t \geq t_0$:

$$\frac{1}{\theta_1\theta_2}\xi\frac{w_0}{w_{t-m}} = \frac{1}{\mu}[\theta_0^{1/f}/(1-p^{1/\pi})^{1/f} - 1] + \xi\frac{w_t}{w_{t-m}}, \qquad t \geq m,$$

$$\frac{1}{\theta_1\theta_2}\xi\frac{w_m}{w_t} = \frac{1}{\mu}[\theta_0^{1/f}/(1-p^{1/\pi})^{1/f} - 1] + \xi\frac{w_m}{w_0}, \qquad t < m.$$

Recalling $\xi(w_x/w_y) \leq e^{-(x-y)/m}$ for $x > y$, we have

$$\xi[1/(w_{t-m}/w_0)] \geq 1/\xi(w_{t-m}/w_0) \geq e^{(t-m)/m};$$

also $\xi(w_t/w_{t-m}) \leq e^{-1}$. We proceed in a similar manner for the $t < m$ case. Either case yields

$$\frac{1}{\theta_1\theta_2}e^{(t-m)/m} \leq \frac{1}{\mu}[\theta_0^{1/f}/(1-p^{1/\pi})^{1/f} - 1] + e^{-1}.$$

Whence substituting $e^{t_0/m} = \theta_1\theta_2$,

$$e^{(t-t_0)/m} \leq [\theta_0^{1/f}/(1-p^{1/\pi})^{1/f} - (1-\mu e^{-1})]e\mu^{-1}, \qquad t \geq t_0,\ \mu \leq 1.$$

We overestimate t, the number of iterations to obtain a probability of termination $> p$ on iteration t, by setting LHS = RHS above. Letting $s = t - t_0$ and $s = s_p$ corresponding to $p = p_t$, we have solving for s_p,

$$s_p = m\ \log\{\theta_0^{1/f}/(1-p^{1/\pi})^{1/f} - (1-\mu e^{-1})\} + m\ \log(e\mu^{-1}), \quad \mu \leq 1.$$

The smallest allowed value of s_p is $s_p = 0$ and this occurs when

$$p = p_{t_0} = (1-\theta_0)^{\pi}.$$

In other words a probability of termination $p \geq (1-\theta_0)^{\pi}$ on some iteration $t_0 + s$ (where s used here is not to be confused with incoming variable P_s),is attained on the "average" for $t \leq t_0 + \gamma s_p$.

Therefore iteration $t = T_p$ corresponding to a probability of termination $p \geq (1-\theta_0)^{\bar{n}}$:

$$T_p = t_0 + \gamma s_p = m \log(\theta_1\theta_2) + \gamma s_p , \quad \gamma = e^{\mu} .$$

For $p > .8^{\bar{n}}$, we can use the very close approximation

$$(1-p^{1/\bar{n}}) \doteq -\log[1-(1-p^{1/\bar{n}})] = -(1/\bar{n})\log p , \qquad .8 \leq p^{1/\bar{n}} \leq 1 .$$

Thus for $p \geq (1-\theta_0)^{\bar{n}}$ and $p > .8^{\bar{n}}$,

$$s_p \doteq m \log(e\,\mu^{-1}) + m \log[\{-\theta_0\bar{n}/\log p\}^{1/f} - (1-\mu e^{-1})] .$$

Suppose $m = 1000, \bar{n} = n-m-1 = 1000, f = 1$ or $f = m/2$, $\theta_0 = .5$, $\theta_1\theta_2 = 4$. How many iterations $t = T_{.01}$ must be performed before the probability of nontermination on iteration t, with $\lambda_s > .14/(m+1)$, is less than $1-p_t = .99$? This means termination is likely within another $1/p_t = 100$ additional iterations. Noting $\mu = .14, \gamma = e^{\mu} = 1.14$, substitution in the above formulas, gives $p > (1-\theta_0)^{\bar{n}}$ and $.8^{\bar{n}}$ so it is okay to use the approximation for $s_{.01}$ which gives $T_{.01} = (1.4 + 7.7\,\gamma)m = 10.1\,m$ if $f = 1$ or $T_{.01} = (1.39 + .017\,\gamma)m = 1.59\,m$ iterations if $f = m/2$.

The inverse function, obtained by solving for $p = p_t$ in terms of $s_p = s$,

$$\Pr(\text{term}) \geq p_t = \{1-\theta_0[1+\mu e^{-1}(-1+e^{s/m})]^{-f}\}^{\bar{n}}, \qquad s = t - t_0 \geq 0 ,$$

will be needed later.

Expected Number of Iterations

If $\overline{p}_t$ is the true probability of termination on step $t, \overline{q}_t = 1 - \overline{p}_t$ and $\overline{Q}_t = \overline{q}_0 \cdot \overline{q}_1 \cdots \overline{q}_t$, then the *expected number of iterations*, by definition, is

$$\begin{aligned} \xi\,\text{ITER} &= 0 \cdot \overline{p}_0 + 1 \cdot \overline{Q}_0 \cdot \overline{p}_1 + 2 \cdot \overline{Q}_0 \overline{p}_2 + \cdots + t\overline{Q}_{t-1}\overline{p}_t + \cdots \\ &= \overline{q}_0 + \overline{q}_0\overline{q}_1 + \overline{q}_0\overline{q}_1\overline{q}_2 + \cdots + \overline{Q}_t + \cdots \quad . \end{aligned}$$

Since $p_t < \overline{p}_t$, the expected iterations beyond $t-1$ is less than

$$Q(t)(1 + q_{t+1} + q_{t+1}q_{t+2} + q_{t+1}q_{t+2}q_{t+3} + \cdots) \quad ,$$

where $q_t = 1 - p_t$ and $Q_t = q_0 \cdot q_1 \cdot \cdots \cdot q_t$. Because $p_t < p_{t+1} < \cdots$, the above is less than

$$Q(t)(1 + q_t + q_t^2 + \cdots) = Q(t)/p(t) , \qquad\qquad Q(t) \leq 1 .$$

It follows that

Lemma: *Expected number of iterations*, $\xi\,\text{ITER} \leq t_0 + \gamma[s_p + (1/p)]$ *where* $\gamma = e^{\mu}$ *is the adjustment so as to include iterations having* $\lambda_s < \lambda_s^* = \mu/(m+1)$ *and* $0 < \mu \leq 1$.

Therefore for any p, $(1-\theta_0)^{\pi} \leq p \leq 1$,

$$\begin{aligned} \xi\text{ITER} \leq\ & m \log(\theta_1\theta_2) + \gamma\, m \log(e\,\mu^{-1}) \\ & + \gamma\, m\{\log[\{\theta_0/(1-p^{1/\pi})\}^{1/f} - (1 - \mu e^{-1})] + 1/pm\} . \end{aligned}$$

A weaker bound can be obtained by dropping the $(1 - \mu\, e^{-1})$ term:

For any $p > (1-\theta_0)^{\pi}$,

$$\xi\text{ITER} \leq m \log(\theta_1\theta_2) + \gamma\, m\{\log(e\,\mu^{-1}) + \frac{1}{f}\log\theta_0 - \frac{1}{f}\log(1-p^{1/\pi}) + 1/pm)\} .$$

At this point we derive some *asymptotic* results. Assume $\bar{n}$ *fixed* as $m \to \infty$. Set $p = (1-\theta_0)^{\bar{n}}$ and note $1/(pm) \to 0$. We have, using the stronger upper bound above

$$\xi\,\text{ITER} \leq m \log(\theta_1\theta_2) \quad \text{for} \quad m \to \infty \quad \text{and} \quad \bar{n} \quad \text{fixed}, \quad \theta_1\theta_2 \geq 1\,.$$

For example $\xi\,\text{ITER} \leq 1.4\,m$ when $\theta_1\theta_2 = 4$. Assume instead $m \to \infty$ and $\bar{n} \to \infty$, then for any fixed p and $\bar{n}$ sufficiently large: $(1-\theta_0)^{\bar{n}} < p$. Again $1/(pm) \to 0$. We can fix p arbitrarily small providing $\bar{n}$ large enough — hence $\log(1 - p^{1/\bar{n}}) \doteq -\log\bar{n} + \log(-\log p) = -\log\bar{n} + b$ where b can be fixed arbitrarily large. Therefore,

$$\xi\,\text{ITER} \leq m \log(\theta_1\theta_2) + \gamma\, m\{\log(e\,\mu^{-1}) + \frac{1}{f}\log\theta_0 + \frac{1}{f}\log\bar{n} - \frac{1}{f}b\}$$

where $1 \leq \gamma \leq e$, $m \to \infty$, $n \to \infty$ and $b > 0$ fixed (arbitrarily large).

Our objective, however, is not to get asymptotic bounds, but bounds for $\xi\,\text{ITER}$. Note any $p > (1-\theta_0)^{\bar{n}}$ can be chosen. Choose $p = e^{-1}$. Since $\gamma \leq e$, the term $1/(pm)$ can be dropped with an error in the bound for $\xi\,\text{ITER} < 2.7\,\gamma < 10$ iterations, actually less since earlier we set $Q(t) = 1$. Note also $-\log\bar{n} \doteq \log(1 - e^{-1/\bar{n}})$. The condition $p > (1-\theta_0)^{\bar{n}}$ becomes $\theta_0\bar{n} \geq 1$. Therefore

$$\begin{aligned}
\xi\,\text{ITER} &\leq m \log(\theta_1\theta_2) + \gamma\, m\{\log(e\,\mu^{-1}) + \frac{1}{f}\log\theta_0 + \frac{1}{f}\log\bar{n}\} \\
&= m[\log(\theta_1\theta_2) + e^{\mu}\{1 + \frac{1}{f}\log\theta_0\bar{n} - \log\mu\}] \\
&= m[\log(\theta_1\theta_2) + E]\,, \qquad\qquad \theta_1\theta_2 \geq 1,\ \theta_0\bar{n} \geq 1\,.
\end{aligned}$$

As a final step we determine μ and $\gamma = e^{\mu}$ so that the bound for E is as small as possible. Note that E is of the form $e^{\mu}(A - \log\mu)$ where $\mu \leq 1$ and

$A = 1 + f^{-1}\log\theta_{\bar\pi} \geq 1$. Setting $dE/d\mu = 0$, we obtain μ as the solution of the equation

$$A = (1 + \mu\log\mu)\mu^{-1}\,.$$

Once μ is obtained as a function of A, we also have

$$E = e^{\mu}\mu^{-1}\,, \qquad \gamma_A = E/A = e^{\mu}/(1 + \mu\log\mu)\,.$$

Corresponding values of $A, \gamma_A, \mu, e^{\mu}, \log\mu$ tabulated below.

$\log\mu$	μ	e^{μ}	A	γ_A
0.0	1.00	2.72	1.00	2.72
-0.5	.61	1.83	1.14	2.65
-1.0	.37	1.44	1.72	2.28
-1.5	.22	1.25	2.98	1.88
-2.0	.14	1.14	5.37	1.57
-3.0	.050	1.05	19.09	1.23
-4.0	.018	1.02	50.60	1.10
$-\infty$	.0000	1.00	$+\infty$	1.00

Therefore, finally,

$$\xi\,\mathrm{ITER} \leq m\log(\theta_1\theta_2) + \gamma_A\, m[1 + \frac{1}{f}\log(\theta_0\bar\pi)]\,,$$

where $\theta_1\theta_2 \geq 1, \theta_0\bar\pi \geq 1$ and $1 \leq \gamma_A \leq e$. γ_A is found in the table above using

$$A = 1 + \frac{1}{f}\log(\theta_0\bar\pi)\,.$$

A better bound for the expected number of iterations can be computed directly from the formula

$$\xi\mathrm{ITER} < t_0 + \gamma(q_{t_0} + q_{t_0}q_{t_0+1} + \cdots + q_{t_0}q_{t_0+1}\cdots q_T)\quad , \qquad q_t = 1 - p_t\,,$$

where the series is truncated at $t = T$ such that $Q_T/p_T < 1$ where $Q_T = q_0 \cdots q_T$. The truncation error is less than $\gamma = e^{\mu}$, $0 < \mu \leq 1$. These bounds on the expected number of iterations have been computed and are given in the tables that follows for various m and n and for $f = 1$ and $f = m/2$ for comparison. Reasonable values for θ_i are $.5 \leq \theta_0 \leq 1$, $1.5 \leq \theta_1\theta_2 \leq 4$. The values of θ_i used are $\theta_0 = 1, \theta_1\theta_2 = 4$. See Table 1 for $f = 1$ and Table 2 for $f = m/2$.

FORMULAE USED TO COMPUTE TABLES

Given $m, n > m+1$ and parameters associated with the distribution from which the columns are drawn:

$$0 < \theta_0 \leq 1, \quad \theta_1\theta_2 \geq 1, \quad f > 0.$$

Determine $\bar{n} = n - m - 1$. If $\theta_0\bar{n} < 1$, set $\theta_0 = 1/\bar{n}$ so that $\theta_0\bar{n} \geq 1$.

To compute μ and γ_A:

$$A = 1 + (1/f)\log_e(\theta_0\bar{n}).$$

Find $\mu \leq 1$:

$$A = (1 + \mu \log_e \mu)/\mu$$

$$\gamma_A = e^{\mu}/(A\mu)$$

Crude Bound $= t_0 + \gamma_A Am + e^{\mu+1}$ where $t_0 = m\log_e(\theta_1\theta_2)$.

Probability of termination on iteration $t \geq p_t$ with $\lambda_s \geq \mu/(m+1)$:

$$p_t = 0, \qquad 0 \leq t < t_0$$

$$p_t = \{1 - \theta_0[1 + \mu e^{-1}(-1 + e^{s/m})]^{-f}\}^{\bar{n}}, \qquad t = t_0 + s \geq t_0$$

$$q_t = 1 - p_t; \qquad Q_t = q_{t_0} \cdot q_{t_0+1} \cdots q_t;$$

Bound $= t_0 + e^{\mu}(Q_{t_0} + Q_{t_0+1} + \cdots + Q_t)\ldots$ terminate when $Q_t < p_t$.

TABLE 1: $f = 1$

BOUND ON EXPECTED NUMBER OF ITERATIONS

AS A MULTIPLE OF THE NUMBER OF EQUATIONS −1

parameter values $f = 1,\ \theta_0 = 1,\ \theta_1\theta_2 = 4$

$m+1 =$ number of equations	$n = 2m$	$n = 2.5m$	$n = 3m$	$n = 3.5m$	$n = 4m$	Crude Bound/m $n = 4m$
	$n =$ Number of Variables					
$m =$ 2	4.8	5.1	5.7	6.2	6.4	8.2
$m =$ 5	5.4	6.2	6.7	7.0	7.3	8.4
$m =$ 10	6.3	6.9	7.3	7.6	7.9	9.0
$m =$ 20	7.0	7.6	8.0	8.2	8.5	9.6
$m =$ 50	7.4	8.4	8.8	9.1	9.3	10.6
$m =$ 100	8.6	9.1	9.4	9.7	9.9	11.3
$m =$ 200	9.3	9.7	10.0	10.3	10.5	12.1
$m =$ 500	10.1	10.6	10.9	11.1	11.4	13.1
$m =$ 1000	10.8	11.2	11.6	11.8	12.0	13.9
$m =$ 2000	11.4	11.9	12.2	12.5	12.7	14.7
$m =$ 5000	12.3	12.8	13.1	13.1	13.3	15.6

TABLE 2: $f = m/2$

BOUND ON EXPECTED NUMBER OF ITERATIONS

AS A MULTIPLE OF THE NUMBER OF EQUATIONS −1

parameter values $f = m/2,\ \theta_0 = 1,\ \theta_1\theta_2 = 4$

$m+1=$ number of equations	$n=$ Number of Variables $n = 2m$	$n = 2.5m$	$n = 3m$	$n = 3.5m$	$n = 4m$	Crude Bound/m $n = 4m$
$m =$ 2	4.8	5.1	5.7	6.2	6.4	8.2
$m =$ 5	4.0	4.5	4.8	5.1	5.2	6.5
$m =$ 10	3.5	3.8	4.0	4.2	4.3	5.6
$m =$ 20	3.0	3.2	3.3	3.4	3.5	5.1
$m =$ 50	2.3	2.4	2.5	2.5	2.6	4.6
$m =$ 100	2.0	2.0	2.1	2.1	2.1	4.4
$m =$ 200	1.7	1.8	1.8	1.8	1.8	4.3
$m =$ 500	1.6	1.7	1.6	1.6	1.6	4.2
$m =$ 1000	1.5	1.5	1.5	1.5	1.5	4.2
$m =$ 2000	1.4	1.4	1.4	1.4	1.4	4.1
$m =$ 5000	1.4	1.4	1.4	1.4	1.4	4.1

Acknowledgments:

I am indebted to Norman Zadeh for his insights which made this paper possible. His intuition as to why the simplex method is efficient, stimulated me to go back to my original formulation to see if it was amenable to analysis. Michael Klass and Michael Taksar gave valuable advice on the randomization model. I am also grateful to Richard Stone who drew my attention to some logical errors in an earlier draft of this paper. Brian Leverich wrote the program for calculating the table of expected number of iterations for various m and n.

Finally a special acknowledgement to Alan Hoffman for his critical review of the analysis and his suggestions for keeping the proofs as free of special assumptions as possible.

References

[1] BORGWARDT, K.H., "Untersuchungen zur Asymptotik der mittleren Schrittzahl von Simplexverfahren in der linearen Optimierung," *Operations-Reseach-Verfahren*, Band *XXVIII*, (1977), pp. 332-345; summary of dissertation.

[2] BORGWARDT, K.H., "Zum Rechenaufwand von Simplexverfahren," *Operations-Reseach-Verfahren*, Band *XXXI*, (1978), pp. 83-97.

[3] DANTZIG, G.B., *Linear Programming and Extensions*, Princeton University Press, 1963, pp. 160-166.

[4] DUNHAM, J.R., D.G. KELLY, AND J.W. TOLLE, "Some Experimental Results Concerning the Expected Number of Pivots for Solving Randomly Generated Linear Programs". TR 77-16 Operations Reseach and Systems Analysis Department, University of North Carolina at Chapel Hill, 1977.

[5] HOFFMAN, A.J., M. MANNOS, D. SOKOLOWSKY and N. WIEGMANN, "Computational Experience in Solving Linear Programs," *J. Soc. Indust. Applied Math.*, Vol. *1*, No. 1 (1953), pp. 17-33.

[6] KLEE, V., "A Class of Linear Programming Problems Requiring a Large Number of Iterations," *Mathematics of the Decision Sciences*, Part 1 (G.B. Dantzig and A.F. Veinott, Jr., eds.), AMS, 1968.

[7] KUHN, H.W. and R.E. QUANDT, "An Experimental Study of the Simplex Method," *Proceeding of Symposia in Applied Mathematics*, Vol. *XV*, AMS (1963).

[8] LIEBLING, T.M., "On The Number of Iterations of the Simplex Method," *Methods of Operations Research, XVII*, V. Oberwolfach-Tagung uber Operations Research, 13-19 August 1977, pp. 248-264.

[9] ORDEN, A., "Computational Investigation and Analysis of Probabilistic Parameters of Convergence of a Simplex Algorithm," *Progress in Operations Research*, Vol. *II*, (A. Prekopa, ed.), North Holland Publishing Company, Amsterdam-London (1976), pp 705-715.

[10] WOLFE, P., CUTLER, L., "Experiments in Linear Programming," *Recent Advances in Mathematical Programming*, R.L. Graves and P. Wolfe, eds., McGraw Hill, 1963, pp. 177-200.

[11] ZADEH, N., "What is the Worst Case Behavior of the Simplex Method," Operations Research Department, Stanford University, draft, (March 1980).

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 14 SOL-80-3R
2. GOVT ACCESSION NO.: AD-A096 130
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): 6 EXPECTED NUMBER OF STEPS OF THE SIMPLEX METHOD FOR A LINEAR PROGRAM WITH A CONVEXITY CONSTRAINT.
5. TYPE OF REPORT & PERIOD COVERED: 9 Technical Report
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): 10 George B. Dantzig
8. CONTRACT OR GRANT NUMBER(s): 15 N00014-75-C-0267, DE-AC03-76SF00326
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Department of Operations Research - SOL, Stanford University, Stanford, CA 94305
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR-047-143
11. CONTROLLING OFFICE NAME AND ADDRESS: Operations Research Program -- ONR, Department of the Navy, 800 N. Quincy St., Arlington, VA 22217
12. REPORT DATE: 11 Oct 80
13. NUMBER OF PAGES: 12 34
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report):

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report):
This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):

18. SUPPLEMENTARY NOTES:

19. KEY WORDS (Continue on reverse side if necessary and identify by block number):
NUMBER OF PIVOTS
EXPECTED ITERATIONS
SIMPLEX METHOD
LINEAR PROGRAMMING
OPERATIONS RESEARCH

20. ABSTRACT (Continue on reverse side if necessary and identify by block number):
Please see attached.

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED 408765 xlt

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

EXPECTED NUMBER OF STEPS OF THE SIMPLEX METHOD FOR A LINEAR PROGRAM WITH A CONVEXITY CONSTRAINT

by

GEORGE B. DANTZIG

Abstract

When there is a convexity constraint, $\Sigma\lambda_j = 1$, each iteration t of the simplex method provides a value z_t for the objective and also a lower bound $z_t - w_t$. The paper studies (1) the expected behavior of (w_t/w_0), (2) probability of termination on the $t-th$ iteration, and (3) the expected number of steps, ξ ITER, when the columns are drawn from a distribution from a three parameter class of distributions. Using estimates based on a random like behavior for covering simplices, it is shown that

$$\xi\, ITER \le m[2.8 + \log_e(\theta_1\theta_2) + (2.8/f)\log_e \bar{n}], \qquad \theta_i \ge 1,$$

where $n = \bar{n} + m + 1$ is the number of non-negative variables, $m+1$ the number of equations. θ_i and f are parameters for varying the distribution.

Reasonable bounds for $1.5 \le \theta_1\theta_2 \le 4$. The critical parameter is $f > 0$. Poor performance can be expected if $f << 1$. For $\theta_1\theta_2 = 4$, and

$$f = 1: \qquad \xi ITER \le 4.2m + 2.8m \log \bar{n},$$

$$f = m/2: \qquad \xi ITER \le 4.2m + 5.5 \log \bar{n}.$$

It is conjectured that $f = m/2$ may be typical of practical problems. If so, for large m and $\bar{n} \le$ some fixed multiple of m, ξ ITER $< 4.2\,m$ iterations as $m \to \infty$. Tighter bounds for $m \le 5000, n \le 4m$ are tabulated. For $m = 1000, n \le 4000$, and $f = m/2$, $\xi ITER < 1.5\,m$.

8

TIC